

To be submitted: *Vision Research*

A Theory of the Benham Top based on Convergent Magno- and Parvo-Cellular Inputs to Cortical Color-Opponent Neurons

Garrett T. Kenyon [1], Dan Hill [1], James Theiler [2], John S. George [1], David W. Marshak [3]

[1] P-21, Biological and Quantum Physics, Los Alamos National Laboratory

[2] NIS-2: Space And Remote Sensing Sciences, Los Alamos National Laboratory

[3] Dept. of Neurobiology and Anatomy, University of Texas Medical School, Houston

To whom correspondence should be addressed:

Garrett Kenyon
P-21, MS D454,
Los Alamos National Laboratory
Los Alamos, NM 87545
phone: 505-667-1900
fax: 505-665-4507
email: gkenyon@lanl.gov

Abstract

The Benham Top (BT) refers to a class of monochromatic patterns that, when rotated, appear colored. We investigated the physiological origins of this phenomenon using a mathematical model of a color-opponent neuron receiving both spectrally-selective parvocellular (P) and luminance-driven magnocellular (M) inputs. For white spots, transient inhibition from the M pathway exactly canceled a transient input from the red-green opponent P pathway, but due to differences in their spatial convergence, this transient input could be unmasked by a BT-like stimulus. The transient response to the BT-like stimulus became weaker as surround sensitivity recovered following saturation, consistent with perceived BT colors. (100/100)

Introduction

The Benham Top (BT) was first described over a hundred years ago when it was reported that a spinning disc imprinted with a characteristic light and dark pattern elicited a perception of multi-colored rings (Benham, 1894). The most common version of the BT is a disk constructed from two filled hemi-circles, one much darker than the other, with the lighter hemi-circle containing several thin, dark arc segments at various radii and azimuthal locations (fig. 1a). When this pattern is rotated at moderate velocities, on the order of 8 Hz (Jarvis, 1977), the concentric rings traced-out by the different arc segments appear colored (fig. 1b). The perceived colors produced by the individual arc segments, also called Benham-Fechner colors, pattern induced flicker colors (PIFCs) or subjective colors, are generally consistent across subjects and lie in the blue/yellow plane (Schramme, 1992; von Campenhausen et al., 1992), indicating that red and green cones contribute equally to the phenomenon. The perception of colors is unchanged when a pair of identical BT stimuli are presented to both eyes simultaneously but at different relative phases, implying that the physiological mechanism, or mechanisms, responsible for the phenomenon are located prior to locus of binocular fusion (for a comprehensive review of BT phenomenology, see von Campenhausen and Schramme, 1995).

Several theories have been advanced to account for Benham-Fechner colors based on differences in the temporal response properties of the three cone types (cf. Courtney and Buchsbaum, 1991). However, timing differences alone cannot be the primary explanation. For the example top shown in fig 1a, the perceived colors associated with the inner- and outer-most arc segments are very different even though the time course of the stimulation is identical. Thus, any model accounting for the BT phenomenon must

incorporate both spatial and temporal processing to be complete. Here, we use a quasi-linear model of a color-opponent neuron in the primary visual cortex to identify a physiological correlate of the BT phenomenon.

The model was constrained by several requirements. 1) We required that flashing black or white spots, regardless of size or temporal frequency, not elicit responses from the cortical color-opponent neuron. This constraint was met by summing equal and opposite red-green opponent inputs. The green input, which was inverted, was delayed by 7 msec, producing a small transient input in the P pathway that was exactly countered by transient inhibition from the M pathway. 2) We required that the model only generate responses to the full BT stimulus and not to any of its component parts, such as the rotating arc segments without the dark hemi-circle or vice-versa. The principle of superposition implies that a purely linear model could not satisfy this constraint. However, by assuming that high-contrast annuli temporarily saturate the M and P surrounds, transient responses to each cycle of a BT-like stimulus could be repeatedly evoked. 3) We required that the model account for the different colors produced by the BT stimulus. When surround sensitivity recovered with a time constant of 100 msec, transient responses to the BT-like stimulus became progressively weaker for annuli representing arc segments with longer onset delays (i.e. located farther from the trailing edge of the dark hemi-circle), thus providing a physiological correlate for the observed differences in perceived color. 4) We required that the model be consistent with the known physiology of early color and luminance processing. Color opponent neurons in the cerebral cortex are likely to receive signals arising from red and green cones via the P pathway (Livingstone and Hubel, 1984; Hubel and Livingstone, 1990; Conway, 2001),

and there is evidence for M pathway input to color processing regions as well (Ferrera et al., 1992; Lachica et al., 1992; Nealey and Maunsell, 1994; Edwards et al., 1995; Ding and Casagrande, 1998; Johnson et al., 2001). Receptive fields in both the M and P pathways possess antagonistic surrounds, but P pathway receptive fields are several times smaller (Derrington and Lennie, 1984; Croner and Kaplan, 1995), implying their convergence onto cortical neurons is much higher. The center-surround spatial organization of P and M pathway cells, along with differences in their spatial convergence, were essential for the model's ability to generate responses to BT-like stimuli.

An alternative model of the BT phenomenon, based upon reports of prominent “rebound” responses in cat X ganglion cells and which, like the present model, considered color opponent inputs with center-surround receptive fields, has previously been proposed (Grunfeld and Spitzer, 1995). However, similar rebound responses do not appear to be elicited in primate P ganglion cells by either small spots or annuli (Benardete and Kaplan, 1997). Our theory of the BT is based on the convergence of M and P pathway inputs onto cortical color-opponent neurons and was tested using simulated light responses that were consistent with published physiological data.

Methods

For simplicity, both the chromatic and luminance input pathways were described by the same impulse kernel, from which the responses to arbitrary patterns of temporal modulation could be calculated by linear convolution. Each input pathway contained both center and surround components, with each component being composed of two input channels, one inverted and delayed relative to the other. In the case of chromatically opponent inputs, the two channels were of complementary color (i.e. red vs. green). The impulse kernel for a single input channel consisted of both forward and feedback components constructed from cascaded low pass filters:

$$G(t) = A \left[\left(\frac{(t/\tau_1)^{n_1} e^{-t/\tau_1}}{n_1!} \right) - R \left(\frac{(t/\tau_2)^{n_2} e^{-t/\tau_2}}{n_2!} \right) \right] \quad (1)$$

with $\tau_1, \tau_2 = 2$ msec, $n_1 = 1, n_2 = 7, R = 0.75$, and A an arbitrary normalization factor determined numerically so that the sustained response along the red input channel to a unit input was equal to one. The above parameters provided a good quantitative fit to the time course of the impulse kernels recorded from the responses of P ganglion cells to center and surround stimulation after the onset latency was subtracted (Benardete and Kaplan, 1997). The above impulse kernel, $G(t)$, was used to calculate the total input to the cortical neuron from the red-selective P pathway centers, denoted by $P_{\text{Red,Center}}$, via linear convolution:

$$P_{\text{Red,Center}}(t) = G(t) \circ [S_{\text{Red,Center}}]_{\text{Max}=1} \quad (2)$$

where $S_{\text{Red,Center}}$ represents the time-dependent input to the red-selective P pathway centers and the notation $[\]_{\text{Max}=1}$ stipulates that the quantity inside the brackets saturates at a value equal to one (i.e. $[x]_{\text{Max}=1} = x, x \leq 1, [x]_{\text{Max}=1} = 1, x > 1$).

The input to the model cortical neuron from the red-selective P pathway surrounds, $P_{\text{Red,Surround}}$, was computed as:

$$P_{\text{Red,Surround}}(t) = \frac{1}{2} \cdot \left(G(t) \circ [0.5 \cdot S_{\text{Red,Center}} + S_{\text{Red,Surround}}]_{\text{Max}=1} + G(t) \circ [0.5 \cdot S_{\text{Green,Center}} + S_{\text{Green,Surround}}]_{\text{Max}=1} \right). \quad (3)$$

The chromatic sensitivity of the P pathway surrounds was assumed to be broadband (i.e. equal contributions from red and green selective channels), regardless of the spectral selectivity of the receptive field center. For simulating the responses to black and white stimuli, such as the BT, the spectral composition of the red and green surrounds was not of fundamental importance, as this aspect of the model only affects the activity evoked by colored stimuli, which was not the primary focus here. Due to the assumed high spatial convergence of P pathway inputs to the model cortical neuron, stimulation of the cortical neuron's receptive field center also activated the P pathway surrounds with the maximum activation being set to 0.5, thus accounting for the dependence of $P_{\text{Red,Surround}}$ on $S_{\text{Red,Center}}$ and $S_{\text{Green,Center}}$. We did not include delays in the activation of the antagonistic surround, which would have complicated the model by introducing additional time constants. The total input to the cortical neuron from the red-selective component of the P pathway, P_{Red} , was then given simply by:

$$P_{\text{Red}}(t) = P_{\text{Red,Center}}(t) - P_{\text{Red,Surround}}(t). \quad (4)$$

For the chromatically opponent green-selective input channel, we used the same impulse kernels as for the red channel with the addition of a small delay,

$$\begin{aligned} P_{\text{Green}}(t) &= [P_{\text{Green,Center}} - P_{\text{Green,surround}}] \\ &= G(t - t_d) \circ [S_{\text{Green,Center}}]_{\text{Max}=1} \\ &\quad - \frac{1}{2} \left(G(t - t_d) \circ [0.5 \cdot S_{\text{Green,Center}} + S_{\text{Green,Surround}}]_{\text{Max}=1} \right. \\ &\quad \left. + G(t - t_d) \circ [0.5 \cdot S_{\text{Red,Center}} + S_{\text{Red,Surround}}]_{\text{Max}=1} \right), \end{aligned} \quad (5)$$

where $t_d = 7$ msec.

The total input from the P pathway to the red-green opponent cortical neuron was simply the sum of the individual inputs from the red- and green-selective components:

$$P(t) = [P_{\text{Red}}(t) - P_{\text{Green}}(t)]. \quad (6)$$

With a trivial exchange of labels, the same model could equally well describe a green-red opponent cortical neuron, and all of the following results would apply. Because the BT activates both red and green channels with equal strength (Schramme, 1992), the implied symmetry between red- and green-selective cortical neurons was essential. We did not compute spatial convolutions in the present model. Instead, the degree to which a given stimulus activated the center relative to the surround was set “by hand”. For a “small”

white spot, we used $S_{\text{Red,Center}}=S_{\text{Green,Center}}=0.5$, $S_{\text{Red,Surround}}=S_{\text{Green,Surround}}=0$, while for a small red spot, we used $S_{\text{Red,Center}}=0.5$, $S_{\text{Green,Center}}=0.25$, $S_{\text{Red,Surround}}=S_{\text{Green,Surround}}=0$. The surround was always equal and opposite in strength to the corresponding center, so that the net input along any channel to a uniform full field stimulus was always zero.

The luminance or M pathway, which by assumption was intrinsically transient, used the same impulse kernels as the P pathway. However, the separate input channels in the luminance pathway were achromatic, which in the present context meant they consisted of equal red and green inputs:

$$\begin{aligned}
 M(t) &= [M_{\text{Center}} - M_{\text{Surround}}] \\
 &= \frac{1}{2} \cdot \left(G(t) \circ [0.5 \cdot S_{\text{Red,Center}} + 0.5 \cdot S_{\text{Green,Center}}]_{\text{Max}=1} \right. \\
 &\quad \left. - G(t - t_d) \circ [0.5 \cdot S_{\text{Red,Center}} + 0.5 \cdot S_{\text{Green,Center}}]_{\text{Max}=1} \right) \\
 &\quad - \frac{1}{2} \cdot \left(G(t) \circ [0.5 \cdot S_{\text{Red,Surround}} + 0.5 \cdot S_{\text{Green,Surround}}]_{\text{Max}=1} \right. \\
 &\quad \left. - G(t - t_d) \circ [0.5 \cdot S_{\text{Red,Surround}} + 0.5 \cdot S_{\text{Green,Surround}}]_{\text{Max}=1} \right).
 \end{aligned} \tag{7}$$

The main difference between the luminance and chromatic pathways was that the surround component of the luminance pathway did not overlap with the cortical neuron's receptive field center and thus was not activated by a “small” spot. We could have allowed the M pathway surrounds to partially overlap the cortical neuron's receptive field center without modifying our basic conclusions, as long as the overlap of the P pathway surrounds remained much greater, but we opted instead for the simplest implementation. The input from the luminance pathway to the cortical opponent neuron was reduced by 50% relative to that of the chromatically opponent pathway. This accounted for the fact that the net P pathway input to the cortical opponent neuron was itself reduced by 50%

due to the extensive overlap of the P pathway centers and surrounds. Equations 1-7 represent a simple model of a cortical opponent neuron that did not respond to flashing black or white spots of any size or temporal frequency, but did respond to flashing red spots and to BT-like stimuli.

By assumption, high-contrast annuli produced maximal activation of the P and M pathway surrounds, $S_{\text{Red,Center}}=S_{\text{Green,Center}}=0.5$, $S_{\text{Red,Surround}}=S_{\text{Green,Surround}}=1$, so that there was no additional surround response to the annulus offset, represented here by the removal of the dark central portion of the annulus so that the remaining stimulus was simply a large uniform spot, $S_{\text{Red,Center}}=S_{\text{Green,Center}}=1$, $S_{\text{Red,Surround}}=S_{\text{Green,Surround}}=1$. Surround sensitivity recovered with a time constant of 100 msec, implemented by removing the saturation condition. Thus, if the annulus was sustained for a period much greater than 100 msec, the offset response became exactly as predicted by the linear model (i.e. the MAX=1 condition was ignored). For intermediate durations, the response to annulus offset was interpolated between these two extremes, approaching that of the linear model with a time constant of 100 msec. The recovery of surround sensitivity was not necessary to unmask responses to BT-like stimuli, but were essential to account for the different colors produced by arc segments with different onset delays.

The responses of the cortical neuron to positive and negative inputs was assumed to be asymmetric, produced by imposing a sharp lower cutoff of -0.1. Such rectification, possibly representing a positive threshold for spike generation, was not necessary to produce temporally modulated responses to BT-like stimuli, but was necessary to unmask a net positive response. The response of the cortical opponent neuron was then given by:

$$C(t) = [P(t) - M(t)]_{\text{Min}=-0.1} . \quad (8)$$

where the notation $[\]_{\text{Min}=-0.1}$ stipulates that the quantity inside the brackets saturates for values less than -0.1 , (i.e. $[x]_{\text{Min}=-0.1} = x, x > -0.1, [x]_{\text{Min}=-0.1} = -0.1, x < -0.1$)

The above model was intended to provide an implementation of the underlying assumptions that was as simple and straightforward as possible. While a more complex model might have been constructed, we believe this would have obscured the basic elements postulated to account for the BT phenomenon.

Results

We constructed a model of a red-selective cortical neuron that received input from both chromatic and luminance channels (fig. 2), identified here with the parvocellular (P) and magnocellular (M) visual pathways, respectively (Hubel and Livingstone, 1990; Merigan and Maunsell, 1990; Merigan et al., 1991; Maunsell, 1992). Because the receptive field center of a cortical red-green color-opponent neuron is much larger than those of typical P pathway neurons (Conway, 2001), we may infer that it receives input from multiple red- and green-selective P pathway cells, with the response polarity of the green input reversed in sign (fig 2b). P pathway neurons have strong antagonistic surrounds (Croner and Kaplan, 1995; Benardete and Kaplan, 1997), which due to their small relative size would overlap extensively within the cortical neuron's receptive field center (fig. 2c). A “small” spot, stimulating half the cortical neuron's receptive field center, would thus activate both the P pathway centers and the P pathway surrounds, effectively reducing the net input from the color-opponent pathway. In contrast, the receptive fields of M pathway neurons are several times larger than those in the P pathway (Derrington and Lennie, 1984; Croner and Kaplan, 1995), and thus the same “small” spot would only weakly activate the surround coming from the luminance channel (fig. 2d). This difference in spatial convergence is critical to the proposed model, which postulates that surround saturation by high-contrast annuli differentially amplifies P vs. M pathway inputs due to differences in their spatial convergence.

A small flashing white spot, which remained on for half a cycle, equal to 128 msec out of a total cycle length of 256 msec, produced a small positive transient input to the cortical opponent neuron from the red–green centers (figs. 3a₁, dashed line)

and a smaller negative transient input from the red–green surrounds (figs. 3a₁, dotted line), yielding a net positive input from the color-opponent pathway (figs. 3a₁₋₂, solid line). The same small flashing white spot produced a small negative transient input from the luminance pathway (figs. 3b₁, solid line), which consisted of a pure center response, as the small spot, by assumption, did not activate the luminance surround. The net inputs from the color-opponent and luminance pathways were equal and opposite, resulting in no net input to the model cortical cell (fig. 3c₁).

A large flashing white spot, here taken to be effectively full field, also produced no net response in the cortical neuron (fig. 3C₂), although in this case the net inputs from the color-opponent (fig. 3a₂, black line) and luminance pathways (fig. 3b₂, black line) were both zero. For a large white spot, the input from the P pathway surrounds (fig. 3a₂, dotted line) exactly cancelled the input from the P pathway centers (fig. 3a₂, dashed line) and similarly the luminance input from the surround channel (fig. 3b₂, dotted line) cancelled the luminance input from the center (fig. 3b₂, dashed line). The above results illustrate how spatially and chromatically opponent inputs were combined so as to ensure that a flashing white spot, regardless of size or temporal frequency, produced no net response in the model cortical neuron.

In contrast to a white spot, the model cortical neuron generated sustained positive responses to red spots flashed at the same temporal frequency (fig. 4c₁). By assumption, a red light activated the green opponent channel at reduced strength, here taken to be 50% of that elicited by a white spot, yielding a sustained positive input from the red–green centers (fig. 4a₁, dashed line). There was also a sustained input from the red–green surrounds of opposite sign (fig. 4a₁, dotted line), but this was not as large as the input

from the red-green centers due to the small spot size used, resulting in a net sustained input from the color-opponent pathway (fig. 4a₁, solid line). In contrast, the luminance channel was only transiently activated by the small red spot (fig. 4b₁, solid line), and at only half the amplitude of that evoked by a small white spot. The model cortical neuron thus therefore exhibiting color selectivity, generating a sustained positive response to a small flashing red spot (fig. 4c₁, solid line) but not to a small white spot (fig. 3c₁, solid line). A large red spot produced an even larger net input to the model cortical cell (figs. 4c₂), due to the increased activation of the red–green center (fig. 4c₂, dashed line), while the achromatic red and green P pathway surrounds exactly cancelled except for a short transient (fig. 4c₂, dotted line). This surround cancellation in the red–green pathway depended on the assumption of broadband spectral sensitivity. Large red spots would have elicited different responses if the red and green surrounds had possessed different spectral sensitivity, but this aspect of the model is not critical for analyzing responses to black and white stimuli, which is our primary interest here.

To mimic a BT-like stimulus, we adopted the following protocol. One cycle of the BT-like stimulus lasted 256 msec, corresponding to a rotation rate of approximately 4 Hz. Each cycle was divided into equal light and dark half-cycles representing the passage of the black and white hemi-circles comprising the standard top. An annulus was used to represent the passage of a thin arc segment through the receptive field center of the cortical neuron. At the beginning of the light half-cycle, the stimulus consisted of an annulus with an effectively infinite outer radius and whose inner portion covered $\frac{1}{2}$ of the cortical neuron's receptive field center. After 32 msec, or $\frac{1}{8}^{\text{th}}$ of the total cycle, the dark central spot was removed, resulting in full-field uniform stimulation. By design, the

annulus produced maximal activation of the red and green surrounds (fig. 5a₁, dotted line) but activated only 50% of the input from the red and green centers (fig. 5a₁, dashed line), producing an initially negative response in the red–green pathway (fig. 5a₁, solid line). Subsequently removing the dark spot at the center of the annulus produced an opposite effect: greater activation of the red and green centers than of the red and green surrounds, yielding a net positive response.

Replacing the annulus by full-field stimulation had relatively little effect on the total activation of the red and green surrounds because these were already saturated by the annulus. When the dark central portion of the annulus was removed, the input from the uncovered red and green centers was thus unopposed by an antagonistic input from the red and green surrounds. Importantly, there was no countervailing amplification in the luminance pathway (fig. 5b₁, solid line), whose centers and surrounds were spatially segregated by assumption. The annulus was therefore able to elicit a net response from the cortical cell (fig. 5c₁, solid line) by saturating the P pathway surrounds, making them insensitive to the subsequent removal of the dark central portion of the annulus. A lower cutoff, possibly representing a threshold potential for spike generation, was used to rectify the output of the cortical neuron and so selectively amplify the positive portion of the response.

A purely linear model that lacked contrast saturation did not respond to a BT-like stimulus (fig. 5c₂), as the activation of the red–green pathway (fig. 5a₂, solid line) became equal and opposite to that of the luminance pathway (fig. 5b₂, solid line). To facilitate direct comparison with the nonlinear model, the linear model was calibrated to produce identical responses to flashing small and large white spots. Responses of the P pathway

centers were unaffected by linearization (compare fig. 5a₂ with fig. 5a₁, dashed lines) as were responses in the luminance pathway (compare fig. 5b₂ with fig. 5b₁). The P pathway surrounds, however, were less strongly activated by the annulus onset and more strongly modulated by the annulus offset after linearization (compare fig. 5a₂ with fig. 5a₁, dotted lines).

To illustrate the critical role played by differences in the spatial convergence of the two input pathways, the red and green surrounds were dissociated from their respective centers and reformed into a distinct spatially-opponent pathway that was restricted to the cortical neuron's receptive field surround, thus making the spatial organization of the red-green opponent pathway identical to that of the luminance pathway. In this situation, the model cortical cell no longer responded to the BT-like stimulus (fig. 5c₃, black line). Indeed, with identical spatial organizations, activation of the red-green pathway (fig. 5a₃) became indistinguishable from that of the luminance pathway (fig. 5a₃), and thus their contributions cancelled exactly. The input to the color opponent pathway was halved in this instance to compensate for the fact that the red and green centers were no longer antagonized by spatially overlapping surrounds. These controls show that the model's responses to the BT-like stimulus depended on two critical non-linear elements: 1) saturation of the red and green surrounds by a high contrast annulus and 2) differences in the spatial convergence of the P vs. M pathways.

The model also exhibited a physiological correlate of the different colors produced by the BT stimulus. Arc segments connected to the trailing edge of the dark hemi-circle appear reddish while those attached to the leading edge appear dark blue or violet (Schramme, 1992). We hypothesize that such differences in perceived color will

be reflected in the degree to which cortical opponent neurons were activated by the different arc segments. To test this hypothesis, the model cortical neuron was driven by a BT-like stimulus in which the delays between the onset of the light half-cycle and that of the annulus were adjusted so as to represent arc segments at different azimuthal locations relative to the trailing edge of the dark hemi-circle (fig 6a). As a function of increasing delay, the net response to the BT-like stimulus progressively declined. The red and green surrounds were immediately saturated upon onset of the light half-cycle, but the surround sensitivity then recovered with a time constant of 100 msec. The longer the delay, the more linear the red and green surrounds became and thus more sensitive to the modulation of a dark central spot representing the onset and offset of the annulus. This linearization, in turn, systematically reduced responses to BT-like stimuli. The persistence of responses at lower levels of surround saturation shows that maximal saturation was not necessary to produce a BT effect. At effectively infinite delays, such as would be produced by eliminating the dark hemi-circle altogether, the red and green surrounds would have fully recovered and thus would react linearly to a flashing dark spot on a light background. The model therefore predicts there would be no activation of a color-opponent cortical neuron produced by the rotating arc segments alone. Rather, the recovery of the surround sensitivity following saturation by a high-contrast annulus predicts that only the full BT-like stimulus will produce a color percept.

Finally, we examined the consequences of varying the effective rotation speed of the BT-like stimulus. Measured as the total integrated output of the cortical neuron over a fixed duration equal to 1 sec, responses to the BT-like stimulus were band-passed, peaking around 8 Hz and falling off at both lower and higher frequencies. Starting at

zero, the initial increase in integrated response with rotation frequency could be attributed to two factors: 1) the extra number of cycles within the fixed integration period, with each cycle producing an approximately equal contribution, and 2) the tendency for surround sensitivity to recover over the duration of a single annulus presentation for very slow rotation speeds. For rotation frequencies above 8 Hz, the decline in the integrated response amplitude follows from the dynamics of the response kernel. At a rotation frequency of 16 Hz, for example, each arc segment lasts for only 8 msec, which was too short to produce large responses in any of the input pathways.

Discussion

We have presented a theory of the Benham Top (BT) based on the responses of a model cortical neuron configured to respond selectively to small red spots. The assumptions incorporated into the model that were critical for producing a physiological correlate of the BT phenomenon are as follows: 1) The model cortical neuron received inputs from both the P and M pathways possessing spatially opponent centers and surrounds. 2) The spatial convergence of the P pathway was much greater than that of the M pathway. 3) There were small transient responses in both the M and P pathways in response to white spots that were equal and opposite, whereas red spots produced sustained responses in only the P pathway. 4) The surrounds of the red-green opponent inputs were temporarily saturated by a high-contrast annulus. 5) Recovery from saturation of the red and green surrounds, reflected by an increase in sensitivity to further light modulation, occurred with a time constant of 100 msec. 6) The input-output properties of the cortical opponent neuron were asymmetric with respect to positive and negative inputs, implemented by imposing a lower cutoff. Below, we describe the supporting evidence for these fundamental assumptions of the model.

Color-selective neurons in the primary visual cortex that receive input from both red and green cones and are chromatically opponent have been previously described (Hubel and Livingstone, 1990; Conway, 2001), consistent with the main features of the model cortical neuron employed here. Color-opponent neurons in the primary visual cortex are thought to receive chromatically selective input from the P pathway (Hubel and Livingstone, 1990; Conway, 2001), but there is evidence for M pathway input to color processing circuits as well (Ferrera et al., 1992; Lachica et al., 1992; Nealey and

Maunsell, 1994; Edwards et al., 1995; Ding and Casagrande, 1998; Johnson et al., 2001). Cells in both the P and M pathways have center-surround receptive field organizations with M pathway receptive fields being several times larger at any given eccentricity (Derrington and Lennie, 1984; Croner and Kaplan, 1995). Thus, the inputs to the model cortical neuron possessed a spatial organization and spatial convergence that was consistent with the known physiology of the M and P pathways.

There is currently insufficient data to determine whether the opponent inputs to color-selective cortical neurons are transiently activated by small white spots. However, such a possibility is not unreasonable. In the red-green opponent pathway, for example, the antagonist green input might arise from an ON-green P-pathway LGN cell whose output is inverted by a cortical interneuron, thus introducing a small delay of a magnitude consistent with that assumed in the model (7 msec). Alternatively, the antagonistic green input might arise from an OFF cell that produced a reduction in tonic excitation at light onset. Since the level of excitation cannot drop below zero, bright stimuli might clip the signals from an OFF input during the response peak, thereby uncovering a small transient from the color-opponent ON input. The M pathway is generally characterized by more transient responses compared to the P pathway (Derrington and Lennie, 1984; Hubel and Livingstone, 1990). While the present model assumed that M pathway inputs consisted solely of transient responses, a small tonic component could probably have been added to the luminance pathway without fundamentally altering our conclusions.

The responses of P pathway cells saturate at high contrast, while those of M pathway cells saturate at lower contrast (Kaplan and Shapley, 1986). The model required that a high-contrast annulus saturate the responses of the red and green surrounds. A

fundamental prediction of the model therefore is that the BT phenomenon should disappear at low contrast, to which P pathway surrounds respond linearly. Furthermore, the model required that the saturation of the red and green surrounds by an annulus be temporary, and that adaptation mechanisms kick in over tens to hundreds of msec to restore the linearity of responses to subsequent modulations. Adaptation mechanisms on somewhat faster time scales have been reported (Yeh et al., 1996). The precise time course of adaptation did not appear to be a critical aspect of the present model, however.

The model can qualitatively account for additional aspects of the BT phenomenon beyond those already described. When the arc segments making up a BT stimulus are very wide, they trace out thick rings that only appear colored along their inner and outer edges while the central portion appears gray (von Campenhausen and Schramme, 1995). In our model, an annulus with a large inner radius, corresponding to a thick arc, would no longer saturate the red and green surrounds and thus would not produce a BT effect. The BT phenomenon also goes away when the rotating pattern is viewed through a pinhole such that only a small part of the pattern is visible (von Campenhausen and Schramme, 1995). Again, this can be accounted for in the present model as resulting from a lack of saturation of the red and green surrounds as the outer radius of the annulus is collapsed inward. The model is also consistent with selective cone demodulation experiments showing that the BT phenomenon requires activation of either red or green photoreceptors, but not blue photoreceptors (von Campenhausen et al., 1992). Clearly, the model of a red-green opponent cortical neuron presented here would not respond at all if both middle and long wavelength cones were demodulated simultaneously. There is, however, evidence that short wavelength cones also provide input to cortical red-green

opponent neurons (Conway, 2001) as well as to cells in the M pathway (Chatterjee and Callaway, 2002). Although the possible contribution from blue cones was not addressed in the present model, the inclusion of such input should not fundamentally effect our main conclusions, especially if it is assumed that blue-yellow cells in the K pathway do not respond to white light. The present model does not account for subjective colors produced by temporal contrast modulation alone, (Festinger et al., 1971), reports of which have been challenged (Jarvis, 1993).

While other quantitative models of the BT phenomenon have been proposed (Courtney and Buchsbaum, 1991; Grunfeld and Spitzer, 1995), the present model is unique in the following respects: 1) the model assumes no fundamental difference in the dynamics of the red and green pathways, except for a small fixed time delay and/or a purely ON/OFF asymmetry, 2) the model employs a linear impulse kernel consistent with the measured phasic-tonic responses of P ganglion cells, and 3) the model utilizes differences in the spatial convergence of the P vs. M pathways. Interestingly, the model implies that the BT phenomenon exists only because color-selective cells in the P pathway possess antagonistic surrounds, and that subjective colors would not be produced if red- and green-selective neurons instead had receptive field structures like those of blue-yellow ganglion cells, which lack conventional surrounds (de Monasterio, 1978; Zrenner and Gouras, 1981). From this perspective, the BT phenomenon is an artifact of using a single visual pathway, the P pathway, to convey information about both chromatic and luminance contrast (Hubel and Livingstone, 1990). Since the P pathway is thought to underlie high spatial acuity, or form, vision (Merigan et al., 1991), antagonistic surrounds are necessary to ensure the transmission of local contrast, as opposed to simple

luminance. Such multiplexing of chromatic and luminance contrast, while efficient, perhaps makes the mammalian color-processing system susceptible to being “falsely” activated by cleverly arranged black and white stimuli. Given the simplicity of the present model, and its strong connection to known physiology and anatomy, we hope that it may provide motivation for future experimental studies of the Benham Top.

Figure Captions

Fig. 1. Example Benham Top (BT). a) Stationary black and white pattern. The BT typically consists of a dark hemi-circle opposite to several short arc segments located at different angular and radial distances. b) Perceived colors. When rotated at angular velocities around 8 Hz, the concentric rings traced out by the different arc segments appear colored. The colors shown are illustrative only and do not attempt to accurately reproduce perceived colors.

Fig. 2. Receptive field structure of a model cortical color-opponent cell receiving convergent input from both P and M pathways. a) Structure of the color-opponent or P pathway input. The model cortical cell received color-opponent input from red- and green-selective channels of opposite polarity. b) High spatial convergence of P pathway receptive field centers. Input from P pathway centers was restricted to the cortical neuron's receptive field center. Mathematically, such inputs were treated as pooled quantities. c) Antagonistic surrounds in the P pathway overlapped extensively. Stimulation of the cortical neuron's receptive field center activated both center and surround components in the P pathway. d) Structure of the luminance or M pathway input. Unlike the color-opponent pathway, the center and surround components of the luminance pathway were spatially segregated.

Fig. 3. Stimulation by flashing white spots. Top illustration: The stimulus was either a small (left column) or large (right column) flashing white spot. a₁) Small spot, color-

opponent pathway. Transient input from the red–green centers (dashed line) was partially countered by an opposite input from the red–green surrounds (dotted line), leaving a reduced net input (solid line). b₁) Small spot, luminance pathway. The M pathway centers were transiently activated by the small spot (dashed line, occluded), yielding a net inhibitory input to the cortical cell from the luminance pathway (solid line). The surround component of the luminance pathway was not activated by the small spot (dotted line). c₁) Small spot, cortical cell output. Inputs from the P and M pathways were equal and opposite, yielding no response in the cortical cell. d₁) Small spot, stimulus protocol. The spot was turned on for one half, or 128 msec, of each stimulus cycle. Height of the stimulus trace was proportional to total center illumination. a₂–d₂) Same interpretation as a₁–d₁, with the small spot replaced by a large spot. a₂–c₂) In both the P and M pathways, the center (dashed lines) and surround (dotted lines) contributions were equal and opposite, so there was no net input to, and thus no net response in, the cortical cell (solid lines). Vertical scale bars are arbitrary and shown for relative comparisons only.

Fig. 4. Stimulation by flashing red spots. Same overall organization as figure 3 but with the white spot replaced with a red spot. a₁) The opponent green channel was only half activated by the red spot, leaving a sustained input from the red–green centers (dashed line) that was only partially countered by the spectrally broad-band red–green surrounds (dotted line), thus leaving a net sustained input from the color-opponent pathway (solid line). b₁) The M pathway centers, which were achromatic, were transiently activated by the red spot at half the amplitude evoked by the white spot (dashed line, occluded),

producing a net transient inhibition of the cortical cell (solid line). Again, the surround component of the luminance pathway was not activated by the small spot (dotted line).

c₁) The sustained input from the P pathway produced a sustained response in the cortical cell. d₁) Stimulus protocol. a₂-d₂) Same interpretation as a₁-d₁, with the small red spot replaced by a large red spot. a₂) The red–green centers (dashed line) were more completely covered, and thus more strongly activated, by the large red spot, whereas the achromatic red–green surrounds (dotted lines) canceled except for a short transient. b₂) Center and surround contributions in the M pathway were equal and opposite. c₂) Larger activation of the P pathway centers produced a larger sustained response in the model cortical cell.

Fig. 5. Model responses to a BT-like stimulus. Top illustration: A BT-like stimulus was constructed from three sequential elements: Left: no spot, Middle: high-contrast black and white annulus, Right: large white spot. *Left column*: Complete non-linear model with overlapping P pathway centers and surrounds. a₁) Color-opponent pathway. Annulus onset produced a large negative input from the red-green surrounds (dotted line) and a smaller positive input from the red-green centers (dashed line). Annulus offset (arrows) produced a second positive response from the red-green centers but no opposing response from the fully saturated red-green surrounds, resulting in a net positive input from the color opponent pathway (solid line). b₁) Luminance pathway. M pathway centers (dashed line) and surrounds (dotted line) resembled P pathway centers and surrounds, producing a net luminance input that was opposite in sign but of smaller magnitude (solid line). c₁) The difference in P vs. M pathway activation produce a temporally modulated

response in the cortical opponent cell, whose integrated area was positive due to an imposed lower cutoff. d₁) Stimulus protocol. The intermediate trace height represents stimulation by the high-contrast annulus. *Middle column*: Linear model with overlapping centers and surrounds. a₂-d₂) Same interpretation as a₁-d₁. In a linear model without saturation, reduced red-green surround activation at annulus onset, and increased activation at annulus offset (arrow), reduced the total P pathway input so that it was equal and opposite to the M pathway input, resulting in no net response in the cortical cell. *Right column*: Non-linear model with spatially segregated P pathway centers and surrounds. a₃-d₃) Same interpretation as a₁-d₁. With center and surround segregation identical in the P and M pathways, their inputs were necessarily equal and opposite despite saturation.

Fig. 6. A physiological correlate of the different colors elicited by the BT. a) Responses of the cortical neuron to BT-like stimuli for different delays of the annulus onset relative to the beginning of the light half-cycle. Red trace: Delay = 0. Yellow trace: delay = 1/8th cycle (32 msec). Green trace: delay = 1/4th cycle (64 msec). Blue trace: delay = 3/8th cycle (96 msec). The magnitude of the temporal modulation decreased with increasing delay. b) Integrated response of the cortical cell to BT-like stimuli with varying temporal frequencies. Different traces correspond to different delays, measured as a fraction of the total cycle length, as described in panel a. The integrated response peaked at around 8 Hz.

References:

- Benardete EA, Kaplan E (1997) The receptive field of the primate P retinal ganglion cell, I: Linear dynamics. *Visual Neuroscience* 14:169-185.
- Benham CE (1894) The artificial spectrum top. *Nature (London)* 51:200.
- Chatterjee S, Callaway EM (2002) S cone contributions to the magnocellular visual pathway in macaque monkey. *Neuron* 35:1135-1146.
- Conway BR (2001) Spatial structure of cone inputs to color cells in alert macaque primary visual cortex (V-1). *J Neurosci* 21:2768-2783.
- Courtney SM, Buchsbaum G (1991) Temporal differences between color pathways within the retina as a possible origin of subjective colors. *Vision Research* 31:1541-1548.
- Croner LJ, Kaplan E (1995) Receptive fields of P and M ganglion cells across the primate retina. *Vision Research* 35:7-24.
- de Monasterio FM (1978) Properties of ganglion cells with atypical receptive-field organization in retina of macaques. *Journal of Neurophysiology* 41:1435-1449.
- Derrington AM, Lennie P (1984) Spatial and temporal contrast sensitivities of neurones in lateral geniculate nucleus of macaque. *J Physiol (Lond)* 357:219-240.
- Ding Y, Casagrande VA (1998) Synaptic and neurochemical characterization of parallel pathways to the cytochrome oxidase blobs of primate visual cortex. *J Comp Neurol* 391:429-443.
- Edwards DP, Purpura KP, Kaplan E (1995) Contrast sensitivity and spatial frequency response of primate cortical neurons in and around the cytochrome oxidase blobs. *Vision Research* 35:1501-1523.
- Ferrera VP, Nealey TA, Maunsell JH (1992) Mixed parvocellular and magnocellular geniculate signals in visual area V4. *Nature* 358:756-761.
- Festinger L, Allyn MR, White CW (1971) The perception of color with achromatic stimulation. *Vision Res* 11:591-612.
- Grunfeld ED, Spitzer H (1995) Spatio-temporal model for subjective colours based on colour coded ganglion cells. *Vision Research* 35:275-283.
- Hubel DH, Livingstone MS (1990) Color and contrast sensitivity in the lateral geniculate body and primary visual cortex of the macaque monkey. *Journal of Neuroscience* 10:2223-2237.
- Jarvis JR (1977) On Fechner-Benham subjective colour. *Vision Res* 17:445-451.
- Jarvis JR (1993) Subjective colour from achromatic flicker. *Vision Res* 33:1131-1132.
- Johnson EN, Hawken MJ, Shapley R (2001) The spatial transformation of color in the primary visual cortex of the macaque monkey. *Nat Neurosci* 4:409-416.
- Kaplan E, Shapley RM (1986) The primate retina contains two types of ganglion cells, with high and low contrast sensitivity. *Proceedings of the National Academy of Sciences of the United States of America* 83:2755-2757.
- Lachica EA, Beck PD, Casagrande VA (1992) Parallel pathways in macaque monkey striate cortex: anatomically defined columns in layer III. *Proc Natl Acad Sci U S A* 89:3566-3570.
- Livingstone MS, Hubel DH (1984) Anatomy and physiology of a color system in the primate visual cortex. *J Neurosci* 4:309-356.

- Maunsell JH (1992) Functional visual streams. *Current Opinion in Neurobiology* 2:506-510.
- Merigan WH, Maunsell JH (1990) Macaque vision after magnocellular lateral geniculate lesions. *Visual Neuroscience* 5:347-352.
- Merigan WH, Katz LM, Maunsell JH (1991) The effects of parvocellular lateral geniculate lesions on the acuity and contrast sensitivity of macaque monkeys. *Journal of Neuroscience* 11:994-1001.
- Nealey TA, Maunsell JH (1994) Magnocellular and parvocellular contributions to the responses of neurons in macaque striate cortex. *J Neurosci* 14:2069-2079.
- Schramme J (1992) Changes in pattern induced flicker colors are mediated by the blue-yellow opponent process. *Vision Res* 32:2129-2134.
- von Campenhausen C, Schramme J (1995) 100 years of Benham's top in colour science. *Perception* 24:695-717.
- von Campenhausen C, Hofstetter K, Schramme J, Tritsch MF (1992) Color induction via non-opponent lateral interactions in the human retina. *Vision Res* 32:913-923.
- Yeh T, Lee BB, Kremers J (1996) The time course of adaptation in macaque retinal ganglion cells. *Vision Research* 36:913-931.
- Zrenner E, Gouras P (1981) Characteristics of the blue sensitive cone mechanism in primate retinal ganglion cells. *Vision Research* 21:1605-1609.